

Although all of the National Education Goals are important, the two that focus on raising student academic achievement — Goals 3 and 5 — are considered by many to be the nation's highest education priorities. Goal 3 calls for all students to demonstrate competency over challenging subject matter, while Goal 5 calls for U.S. students to be first in the world in mathematics and science achievement.

State policymakers need good information to help them monitor their state's progress toward these Goals. First, policymakers need to know whether student achievement is increasing over time, so that they can determine whether educational programs and policies are having the desired effect. Second, policymakers need to be able to benchmark their state against other states and countries to see how their students' academic performance compares to the best in the nation and the best in the world. Third, policymakers need to know how different groups of students are performing academically, so that they can target educational services appropriately.

This report provides all three types of information. Its purpose is to summarize the amount of progress that each state has made in raising student academic achievement in mathematics and science since the National Education Goals were established in 1990.

Report format

This report contains four pages of information for the United States, each state, the District of Columbia, and five U.S. territories.¹ The first three pages in each set measure progress toward Goal 3, using student achievement data from the National Assessment of Educational Progress (NAEP). NAEP was authorized by Congress in 1969, and is the only nationally representative and ongoing assessment that measures what students know and are able to do in different subject areas. Congress expanded NAEP to

allow the reporting of comparable state-by-state results, beginning with the 1990 mathematics assessment. Participation in state-level NAEP is voluntary, and has increased from 40 states and territories in 1990 to 45 in 1996.

NAEP results are reported for the United States and for each participating state in mathematics (Grades 4 and 8) and science (Grade 8). Thus far, these are the only grades in which NAEP mathematics and science assessments have been administered at the state level. Since 1990, mathematics has been assessed twice at Grade 4 (in 1992 and 1996) and three times at Grade 8 (in 1990, 1992, and 1996). Science has been assessed once (in 1996).²

Each of the state pages in this report shows:

- how much progress the state has made over time;
- how the state's latest academic performance compares to that of the United States and other states; and
- how different subgroups of students in the state performed on the most recent NAEP assessment.

Gold stars are awarded to states that have shown a significant increase in the percentage of students in their state who meet the National Education Goals Panel's performance standard.³ The Goals Panel's performance standard is based on three achievement levels set by the National Assessment Governing Board to describe the quality of student achievement on NAEP: Basic, Proficient, and Advanced. The Basic level represents partial mastery of necessary knowledge and skills; the Proficient level represents solid academic performance; and the Advanced level represents superior performance.⁴ The Goals Panel has set its performance standard at the Proficient or Advanced levels on NAEP. The Goals Panel considers performance at these two highest levels as evidence that students have demonstrated competency over challenging subject matter.

¹ The term "state" is used hereafter in this report to refer to the 50 states, the District of Columbia, and the territories.

² See Appendix B for national and state-level NAEP administration schedules.

³ In this report, "significance" refers to statistical significance and indicates that the observed differences are not likely to have occurred by chance. All differences in this report that are termed "statistically significant" are measured at the 0.05 level.

⁴ Bourque, M.L., Champagne, A.B., & Crissman, S. (1997, October). *1996 Science performance standards: Achievement results for the nation and the states*. Washington, DC: National Assessment Governing Board.

The fourth page in each set of state pages in this report shows how close each state is to achieving Goal 5. Although Goal 5 calls for the United States — not each individual state — to be first in the world in mathematics and science, the majority of states must be at world-class levels of performance in mathematics and science if we expect the nation to attain first-in-the-world status. International comparisons of student achievement in 8th grade mathematics and science are presented, using data from a newly released research study.⁵ This study statistically links state results from the 1996 NAEP with country results from the 1995 Third International Mathematics and Science Study (TIMSS). TIMSS is the most comprehensive international study of mathematics and science achievement conducted to date. TIMSS tested half a million students in 41 countries in 30 different languages. Participating countries included the United States, as well as some of the United States' chief economic competitors and trading partners, such as Japan, Germany, Canada, Korea, Singapore, and Hong Kong.

Linking the two assessments allows us to predict how each state would have performed on TIMSS, relative to the 41 countries that actually participated in the international assessment, on the basis of each state's NAEP performance. The authors of the linking study caution that the technique used to link the two tests can provide only limited information, since NAEP and TIMSS cover different content and were taken by different groups of students at different times. Nevertheless, the technique can provide broad comparisons that tell states which countries' students would be expected to score significantly higher than, similar to, or significantly lower than their own students in mathematics and science on this international assessment. In this report, gold stars signifying "world-class performance" are awarded to those states that would be expected to score as well as, or better than, 40 or more of the 41 participating TIMSS nations in mathematics or science.

Value to states

This report shows three of the ways in which NAEP data can be a valuable source of information for states:

1. NAEP can be used to monitor educational progress over time.

One of the most common uses of NAEP is to monitor trends in academic performance to see whether student achievement is improving over time. This is possible because NAEP is designed to repeat assessments in each subject area at least every four years. This feature enables policymakers to answer questions such as: Has student performance improved since my state established new statewide standards in science? Are more 8th graders in my state considered Proficient in mathematics since my state began requiring all 8th graders to take algebra?

Improvement Over Time is presented in Part 1 on the first three pages for each state in this report, beginning on p. 12. The percentages of students who scored at or above the Proficient level on NAEP mathematics and science assessments will be tracked over a ten-year period, from the establishment of the National Education Goals in 1990, until the year 2000.

2. NAEP can be used to benchmark state performance against the best in the nation and the best in the world.

Because NAEP scores are comparable across states, policymakers can use NAEP to answer questions such as: How does my state compare to neighboring states or to the highest-performing states in the country?⁶ It is also possible to use NAEP scores in a more limited way to predict relative performance on a related assessment such as TIMSS, so that states can benchmark their performance against top-performing nations in mathematics and science. Policymakers can use results from the NAEP/TIMSS linking study to answer questions such as: How many nations would be expected to outperform my state in 8th grade mathematics? How would my

⁵ Johnson, E.G., & Siegendorf, A. (1998, May). *Linking the National Assessment of Educational Progress and the Third International Mathematics and Science Study: Eighth grade results*. Report prepared for the U.S. Department of Education, National Center for Education Statistics, NCES 98-500. Washington, DC: U.S. Government Printing Office.

⁶ Although NAEP scores are comparable, the reader should bear in mind that many variables of interest to state policymakers can contribute to differences in state performance, such as available resources, curricula, educational practices, etc. The results presented in this report do not control for these variables.

state be expected to perform in comparison to the United States' major trading partners in 8th grade science?

State Comparisons are presented in Part 2 on the first three pages for each state in this report, beginning on p. 12. Each state's performance is compared to the nation and to other states on the most recent NAEP mathematics and science assessments.

International Comparisons are presented on the fourth page of each set of state pages. Each state's predicted performance on TIMSS is compared to the actual performance of the 41 participating TIMSS nations.⁷ Countries are clustered in alphabetical order in three groups: those that would be expected to perform significantly higher than, significantly lower than, or not significantly different from the particular state in 8th grade mathematics and science.

3. NAEP can be used to monitor whether all groups of students in a state are achieving at high levels.

Goal 3 specifies that all students will demonstrate competency over challenging subject matter. Because NAEP data can be broken out by subgroups, policymakers can use NAEP to answer questions such as: Are similar proportions of boys and girls in my state considered Proficient in mathematics and science? Do minority students score as well as White students? Do large achievement gaps exist between urban and non-urban students?

Subgroup Performance is presented in Part 3 on the first three pages for each state in this report, beginning on p. 12. This section shows how many students in different subgroups scored at or above the Proficient level on the most recent NAEP mathematics and science assessments. Results are presented by sex, race/ethnicity, parents' highest level of education, school location, and eligibility for free/reduced-price lunch programs.

Interpreting the results

NAEP is a large-scale assessment intended for monitoring trends in student performance and is not administered to every student. Instead, samples of students are selected to take the test. This enables states to use smaller, cost-efficient samples to predict how the entire student population would have performed on an assessment without testing all of them. This is similar to a public opinion poll that predicts, with a certain degree of confidence, how all individuals would have responded to a set of questions had they all been polled.

It is important to note that any estimate based on a sample, whether it is from a NAEP assessment or a public opinion poll, contains a small amount of error. The estimate would be slightly higher or slightly lower if a different sample were chosen. Public opinion polls account for this error when they caution that their results are "accurate within plus or minus two percentage points." In the same way, we must account for the uncertainty in NAEP results, whether we are comparing progress over time, performance among states, or performance among subgroups of students within a state.

We account for the uncertainty by using a formula to calculate a standard error for each estimate.⁸ The standard error tells us how precise the estimate is. The closer the standard error is to zero, the more precise the estimate. Although sample size is only one of several factors that influence the size of the standard error, as a general rule, larger samples yield more precise estimates and smaller standard errors.

If we want to examine differences between groups — for example, to determine whether one state performed at a higher level than another did — we must apply a statistical test to tell us whether there are likely to be differences in actual performance between groups in the entire population.⁹ The statistical test takes into account the size of the standard errors for each group's score, as well as the difference between the scores. If the test indicates that the groups

⁷ In 1995, a representative sample of 8th graders in Minnesota took the same mathematics and science assessments as the students in the 41 participating TIMSS nations. Results shown for Minnesota, therefore, are based on actual scores, not estimated scores. Missouri and Oregon also took the same TIMSS assessments in 1997, but their results have not yet been publicly released.

⁸ See Appendix A for formulas and more detailed technical information. See Appendix C for tables of standard errors.

⁹ See Appendix A for a discussion of the statistical procedures used to control the amount of error introduced when multiple comparisons are made. For more detailed information, see also the Technical Report of the NAEP 1996 State Assessment Program in Mathematics.

in the entire population are likely to perform differently, we say that the difference is statistically significant. This means that the differences are not likely to have occurred by chance — we can be confident that performance has changed over time or one group has outperformed another.

This should be kept in mind when reviewing the data on the state pages that follow. In Part 1, for example, it may appear that the percentage of students who scored at the Proficient level or higher on NAEP has gone up over time, but the change is reported as “not significant.” This occurs because even though there is a difference in scores, it is not statistically different. Because each percentage is an estimate which has some uncertainty associated with it, it is possible for a small gain to be significant in one case, while a larger percentage-point gain can fail to be significant in another.

The same caution must be exercised when interpreting the results in Parts 2 and 3. In Part 2, it would not be accurate to rank individual states strictly by the percentages of students who scored at or above Proficient. Instead of ranking individual states, it is more useful to talk about states’ performance in terms of clusters of states that performed significantly higher than, significantly lower than, or not significantly different from a particular state. On p. 21, for example, the percentage at or above Proficient in 8th grade mathematics for Alaska was 30% in 1996, while Colorado was 25% and Maryland was 24%. When accounting for error, however, Maryland (but not Colorado) is judged to have a similar achievement level to Alaska, even though the percentage for Colorado was larger than Maryland’s.

Similarly, in Part 3, it would not be accurate to conclude that one subgroup of students outperformed another based solely on the percentages listed on the graph. An observed difference of 3 percentage points between males and females, for example, may not be statistically significant when standard errors are taken into account. In order to keep the graphs in Part 3 as clear and as

readable as possible, we have not attempted to flag subgroup differences on the graphs themselves. Instead, statistically significant differences between subgroups are summarized in Appendix D.

Finally, readers should use caution when interpreting the results of the NAEP/TIMSS linking study in this report. The purpose of the linking study is to compare states to nations, not states to states or nations to nations. State-to-state comparisons, using comparable NAEP data, appear in Part 2 on the first three pages for each state. Nation-to-nation comparisons, using comparable TIMSS data, appear on the international comparisons page for the United States (see p. 15). Because the results of the NAEP/TIMSS linking study can offer only approximate comparisons of performance of individual states relative to the 41 participating TIMSS nations, nations and states are simply listed in alphabetical order and actual scores are not shown.¹⁰

Findings — Improvement Over Time

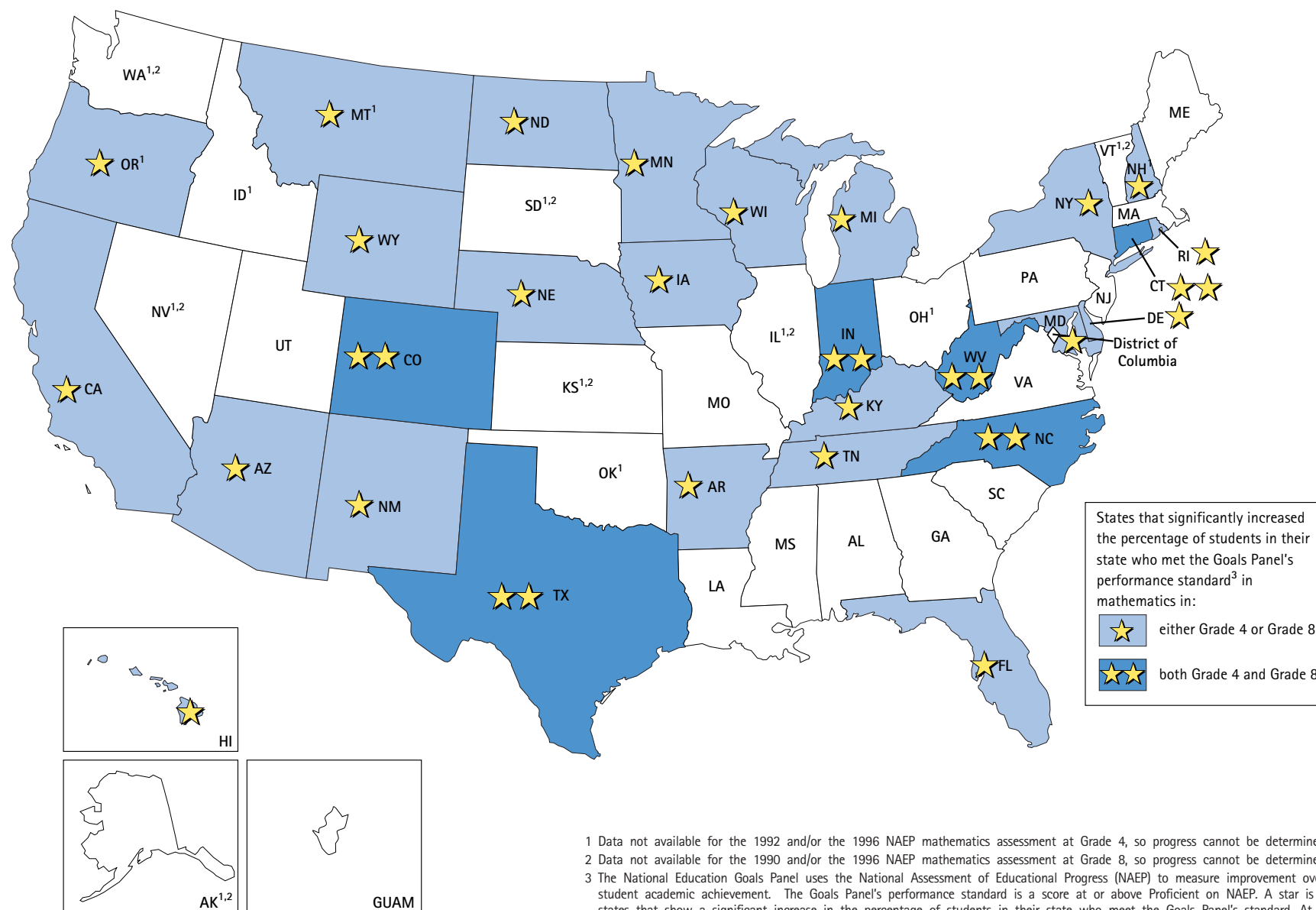
The percentage of students who met the Goals Panel’s performance standard (that is, a score at or above Proficient on NAEP) increased significantly during the 1990s:

- nationally and in 7 states in 4th grade mathematics; and
- nationally and in 27 states in 8th grade mathematics.

In no state has achievement declined by an amount that is statistically significant.

The 28 states that earned stars for improvement over time are shown on the map in Figure 1. At present, the maximum number of stars that a state can earn for improvement in student academic achievement is two (in 4th grade mathematics and in 8th grade mathematics). A star for improvement cannot yet be earned in 8th grade science, because NAEP has assessed science only once at the state level.

¹⁰ For more detailed technical information about the NAEP/TIMSS linking study, see the forthcoming report from the National Center for Education Statistics, *Linking the National Assessment of Educational Progress and the Third International Mathematics and Science Study at the eighth grade: A research report*.



1 Data not available for the 1992 and/or the 1996 NAEP mathematics assessment at Grade 4, so progress cannot be determined.

2 Data not available for the 1990 and/or the 1996 NAEP mathematics assessment at Grade 8, so progress cannot be determined.

3 The National Education Goals Panel uses the National Assessment of Educational Progress (NAEP) to measure improvement over time in student academic achievement. The Goals Panel's performance standard is a score at or above Proficient on NAEP. A star is awarded to states that show a significant increase in the percentage of students in their state who meet the Goals Panel's standard. At present, the maximum number of stars that a state can earn for improvement over time is two (in 4th grade mathematics and 8th grade mathematics). A star for improvement cannot yet be earned in 8th grade science, because NAEP has assessed science only once at the state level.



Twenty-eight states have earned at least one star (out of a possible two) for improvement over time in mathematics in either Grade 4 or Grade 8:

- | | | |
|----------------|--------------------|-------------------|
| 1. Arizona | 11. Kentucky | 21. North Dakota |
| 2. Arkansas | 12. Maryland | 22. Oregon |
| 3. California | 13. Michigan | 23. Rhode Island |
| 4. Colorado | 14. Minnesota | 24. Tennessee |
| 5. Connecticut | 15. Montana | 25. Texas |
| 6. Delaware | 16. Nebraska | 26. West Virginia |
| 7. Florida | 17. New Hampshire | 27. Wisconsin |
| 8. Hawaii | 18. New Mexico | 28. Wyoming |
| 9. Indiana | 19. New York | |
| 10. Iowa | 20. North Carolina | |



Six of these states have earned two stars (out of a possible two) for improvement over time in mathematics in both Grade 4 and Grade 8:

1. Colorado
2. Connecticut
3. Indiana
4. North Carolina
5. Texas
6. West Virginia

- *Seven states that made significant gains in mathematics were also among the highest-performing states* in the nation:*

1. Connecticut
2. Iowa
3. Minnesota
4. Montana
5. Nebraska
6. North Dakota
7. Wisconsin

Findings — State Comparisons

Mathematics — Grade 4

National Performance

In 1996, 21% of U.S. 4th graders in public and nonpublic schools scored at the Proficient level or higher on the NAEP mathematics assessment.

State Performance

In 1996, the percentage of public school 4th graders who scored at the Proficient level or higher on the NAEP mathematics assessment ranged from 3% in the lowest-performing states to 31% in the highest-performing states.

Highest-performing* states

Connecticut	31%
Minnesota	29%
Maine	27%
Wisconsin	27%

Mathematics — Grade 8

National Performance

In 1996, 24% of U.S. 8th graders in public and nonpublic schools scored at the Proficient level or higher on the NAEP mathematics assessment.

State Performance

In 1996, the percentage of public school 8th graders who scored at the Proficient level or higher on the NAEP mathematics assessment ranged from 5% in the lowest-performing states to 34% in the highest-performing states.

Highest-performing* states

Minnesota	34%
North Dakota	33%
Montana	32%
Wisconsin	32%
Connecticut	31%
Iowa	31%
Maine	31%
Nebraska	31%
Alaska	30%

* Highest-performing states are defined as those in which the percentage of students who scored at or above Proficient on NAEP was significantly higher than the percentage of students who did so nationally.

Science — Grade 8

National Performance

In 1996, 29% of U.S. 8th graders in public and nonpublic schools scored at the Proficient level or higher on the NAEP science assessment.

State Performance

In 1996, the percentage of public school 8th graders who scored at the Proficient level or higher on the NAEP science assessment ranged from 5% in the lowest-performing states to 41% in the highest-performing states.

Highest-performing states*

Maine	41%
Montana	41%
North Dakota	41%
Wisconsin	39%
Massachusetts	37%
Minnesota	37%
Connecticut	36%
Iowa	36%
Nebraska	35%
Wyoming	34%

and Asian/Pacific Islander students who scored at the Proficient level or higher on NAEP.

- However, in the majority of cases at both the national and state levels, the percentages of White students who scored at the Proficient level or higher were significantly greater than the percentages of American Indian/Alaskan Native, Black, and Hispanic students who met this standard. This was true for 4th grade mathematics, 8th grade mathematics, and 8th grade science.

Differences by Parents' Highest Level of Education

- Nationally and in almost every case at the state level, students whose parents had some education beyond high school or whose parents were college graduates outperformed students who reported that neither of their parents had graduated from high school.

Differences by School Location

- At the national level, students who attended school in urban fringes/large towns outperformed those who attended school in central cities in 4th grade mathematics and 8th grade mathematics. This was also true in roughly one-third of the states.

Differences by Poverty

(as measured by eligibility for free/reduced-price lunch program)

- In all cases — nationally and in every state — students who were not eligible for the free/reduced-price lunch program outperformed students who were eligible for this program. This was true across all subjects and grades.

Findings — International Comparisons

In 8th grade mathematics, the United States scored higher than 7 countries, lower than 20, and not significantly different from 13. In 8th grade science, the United States scored higher than 15 countries, lower than 9, and not significantly different from 16.

Findings — Subgroup Performance¹¹

Differences by Sex

- Nationally and in 9 out of 45 states, the percentage of male students who scored at or above Proficient in 4th grade mathematics was higher than the percentage of females who did so.
- In 6 out of 43 states, males outperformed females in 8th grade mathematics. There was no significant difference at the national level.
- In 19 out of 42 states, males outperformed females in 8th grade science. There was no significant difference at the national level.

Differences by Race/Ethnicity

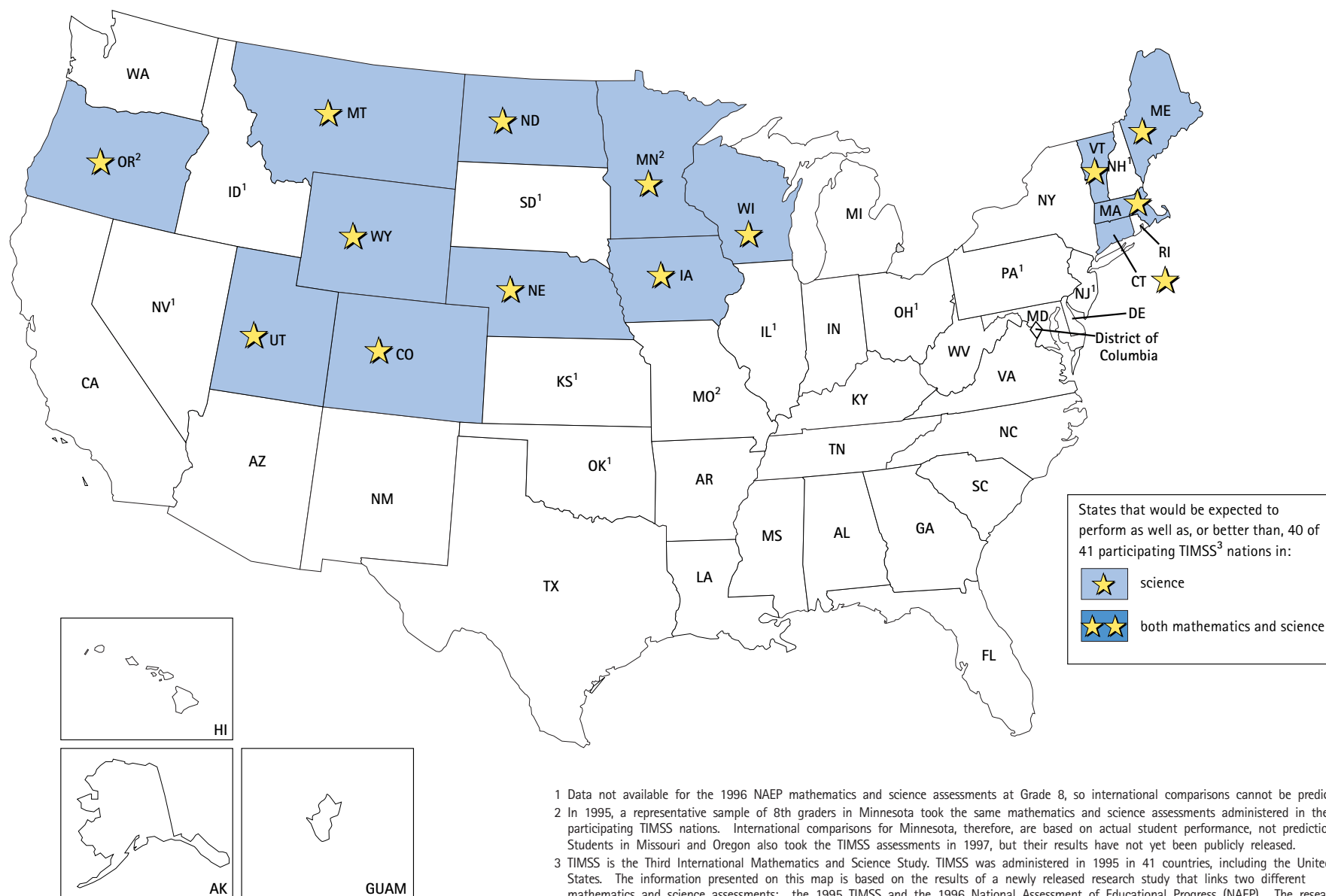
- At the national level and in most of the states, there were no significant differences between the percentages of White

* Highest-performing states are defined as those in which the percentage of students who scored at or above Proficient on NAEP was significantly higher than the percentage of students who did so nationally.

¹¹ The reader is cautioned to avoid interpreting subgroup differences in this section of the report and in Appendix D as causal relationships.

Figure 2

State Progress Toward Goal 5: World-Class Performance in Mathematics and Science Achievement



- 1 Data not available for the 1996 NAEP mathematics and science assessments at Grade 8, so international comparisons cannot be predicted.
- 2 In 1995, a representative sample of 8th graders in Minnesota took the same mathematics and science assessments administered in the 41 participating TIMSS nations. International comparisons for Minnesota, therefore, are based on actual student performance, not predictions. Students in Missouri and Oregon also took the TIMSS assessments in 1997, but their results have not yet been publicly released.
- 3 TIMSS is the Third International Mathematics and Science Study. TIMSS was administered in 1995 in 41 countries, including the United States. The information presented on this map is based on the results of a newly released research study that links two different mathematics and science assessments: the 1995 TIMSS and the 1996 National Assessment of Educational Progress (NAEP). The research study was designed to compare states to nations, and predicts how individual states would have performed on TIMSS, given their NAEP scores. At present, the maximum number of stars that a state can earn for world-class academic performance is two (in Grade 8 mathematics and Grade 8 science), although no state has earned a star in mathematics. For more information, see Appendix A.

When compared to our chief economic partners, the United States is in the bottom half in mathematics and around the middle in science.

The expected performance of individual states on the TIMSS mathematics and science assessments varied widely. In mathematics, the number of countries that would be expected to outperform a given state ranged from 6 to 38. In science, the number ranged from 1 to 38.

States that earned gold stars for "world-class performance" in mathematics and science are shown on the map in Figure 2. Stars were awarded to states that would be expected to score as well as, or better than, 40 or more of the 41 participating TIMSS nations in mathematics or science. The maximum number of stars that a state can earn for world-class performance is two (one in mathematics and one in science), although no state earned a star in mathematics.



In science, 14 states earned a star for world-class performance. Students in only one nation — Singapore — would be expected to outperform the 8th graders in these states in science:

- | | |
|------------------|-----------------|
| 1. Colorado | 8. Nebraska |
| 2. Connecticut | 9. North Dakota |
| 3. Iowa | 10. Oregon |
| 4. Maine | 11. Utah |
| 5. Massachusetts | 12. Vermont |
| 6. Minnesota | 13. Wisconsin |
| 7. Montana | 14. Wyoming |


Conclusions

Are states making progress toward Goal 3 of the National Education Goals by increasing student achievement in mathematics

and science? We cannot answer this question for science yet because NAEP has assessed science only once at the state level. However, in mathematics the answer is "yes." The majority of states that participated in NAEP assessments during the 1990s have shown significant improvements in student academic achievement in mathematics in at least one grade. Twenty-eight states have earned at least one star (out of a possible two) for improvement over time, and six states have earned two. From this perspective, the majority of states have moved closer to the Goal in mathematics.

How close are states to achieving the world-class levels of performance in mathematics and science indicated in Goal 5? Results of the NAEP/TIMSS linking study suggest that no state would likely place first in the world in mathematics. States with the highest NAEP scores in the nation would be expected to trail at least six countries if they were to take the TIMSS assessment. However, 14 states would be expected to achieve world-class levels of performance in 8th grade science. Of the 41 nations that participated in TIMSS, only Singapore would be expected to score significantly higher than the 8th graders in these states.

The challenge before us now is to keep this momentum going — to accelerate student academic progress in more states, in more subject areas, in more grades, and among all students. The National Education Goals Panel will continue to monitor state progress as new state-level NAEP assessments are administered in reading, writing, mathematics, and science between now and the end of the decade, and in 1999, when the international TIMSS mathematics and science assessments are scheduled to be repeated in approximately 40 countries. Future Goals Panel reports will describe educational programs and policies implemented by states that have made significant progress in raising student academic achievement. This information will be available on the Goals Panel's Web site, www.negp.gov, as part of a series of reports on promising state practices.



The National Education Goals Panel remains convinced that states want and need good information that will help them gauge the success of their education improvement efforts. This document shows three ways in which NAEP and TIMSS data can help state policymakers measure their state's progress toward Goals 3 and 5:

- by monitoring educational progress over time;
- by benchmarking their students' academic performance against the best in the nation and the best in the world; and
- by monitoring the extent to which all groups of students in their state are achieving at high levels.

The Goals Panel strongly encourages states to continue participating in NAEP assessments and to consider participating in the next administration of TIMSS, so that policymakers and the public can determine whether educational programs and policies are producing the desired results — students who are competent, knowledgeable, and capable.